



Journal of Statistical Software

April 2009, Volume 30, Book Review 2.

<http://www.jstatsoft.org/>

Reviewer: Brigid Wilson
University of California, Los Angeles

A Handbook of Statistical Analyses Using SAS (Third Edition)

Geoff Der and Brian S. Everitt
Chapman & Hall/CRC, Boca Raton, FL, 2009.
ISBN 978-1-58488-784-3. 391 pp. USD 54.95.
<http://support.sas.com/HandbookofStatisticalAnalyses>

A Handbook of Statistical Analyses Using SAS (Third Edition) by Geoff Der and Brian S. Everitt aims to be a self-contained volume to assist in data analysis and to introduce the reader to the options available with SAS. Given the growing number of procedures available in SAS, this aim becomes increasingly ambitious with each edition. The authors largely succeed in the content of the book, expanding on the second edition of the book with practical examples of modeling and graphing techniques that have been added to SAS in recent years. They cover a wide range of topics efficiently – the explanations are brief, but not overly simplistic; the examples are sufficient and never excessive. Their efforts are slightly undermined by some organizational shortcomings, but this a good resource for the user who is acquainted with the very basics of SAS, but unsure of how to conduct analysis. The first chapter, an introduction to SAS, presents the interface and the SAS language's DATA and PROC steps, as well as the Output Delivery System (ODS) and concludes with two indispensable pages on preventing and correcting errors in SAS programming. In this introduction, the authors do not attempt to exhaustively demonstrate the options available, but rather showcase a few that are used most often. They employ additional options throughout the book in examples. In each such case, the option used and the reason why it is appropriate to the example are explained.

After the introduction, each chapter describes a method of analysis. Typically, a chapter begins with the introduction of several example datasets and related questions that motivate the featured method. The authors use these data sets to provide examples of data read-in and data management strategies. They then present the theory behind the method and devote a section of the chapter to how it is carried out in SAS. Before presenting code for a model, the authors present ways of looking at the data with tables and plots that illustrate what is to come in the models. With this gradual approach, the authors illustrate good practices for catching data errors or outliers and preventing poor model choices. Most chapters include at least two examples of the SAS procedure(s) used to carry out the given method with different options in each example. This shows the reader some of the options that can be specified and the syntax needed to make these specifications. Interpretations of the analysis examples are brief, and the authors assume that the reader knows where in the output to look for the

information relevant to the analysis questions. Each chapter ends with a series of exercises that reinforce the deliberate steps towards a model. An appendix provides the answers to selected exercises. The reader can download all of the described data sets and code in the book to complete these exercises.

The analysis topics covered begin with simple inference and then move to analysis of variance (ANOVA) and ordinary-least-squares regression. The authors then address more complex modeling techniques including logistic regression, generalized linear models, and generalized additive models. Since the publication of the second edition in 2002, the releases of SAS 9.1 and SAS 9.2 have expanded the software's capabilities in these areas. The new edition includes examples of conditional logistic regression and the use of PROC GAM for additive models with loess and spline fits. The authors next present methods for analyzing repeated measures, longitudinal, and survival data. The longitudinal data section includes a chapter on generalized estimating equations and generalized mixed models that is new to this edition. This chapter focuses on PROC GLIMMIX and the updated PROC GENMOD options for correlated observations. The last four chapters illustrate multivariate techniques in SAS. Most plots in the book are generated using the SGPLOT procedure introduced in SAS 9.2 or the ODS graphics available in certain procedures and demonstrate the improvements in SAS graphics over the past decade.

The chapters range in length from ten to twenty-five pages. Most could be read independently of the rest of the text to provide the reader with a quick overview of a topic and a set of examples with code. A few topics (ANOVA, longitudinal data) are addressed over multiple chapters. The survival analysis chapter reads as though there is enough material for two chapters and the concepts in the chapter might have been better illustrated if examples followed explanations more immediately.

The greatest failing of the book may be the index. It has surprisingly few listings under each letter heading, but the listings tend to have multiple sublistings. The index listings do not include any SAS commands and exclude some common phrases like "regression"; instead, analysis techniques often appear as sublistings below the names of the example data sets in which they were demonstrated. For instance, "Poisson regression" is not listed under "P" or "R", but under "A" for "Australian school attendance". This organizing scheme defeats the purpose of an index. This may seem a trivial drawback, but it greatly hinders quick and easy reference. One assumes these are two aims of a self-described "handbook". It should be noted that the index of the second edition does not share these flaws and is quite comprehensive.

Reviewer:

Brigid Wilson
University of California, Los Angeles
Department of Statistics
Los Angeles, CA 90024, United States of America
E-mail: brigid@stat.ucla.edu
URL: <http://www.stat.ucla.edu/~brigid/>

Journal of Statistical Software
published by the American Statistical Association
Volume 30, Book Review 2
April 2009

<http://www.jstatsoft.org/>
<http://www.amstat.org/>

Published: 2009-04-27
